DOCUMENT RESUME

ED 105 717 FL 006 293

AUTHOR Moravcsik, Edith A.
TITLE Determination. Working Papers on Language Universals, No. 1.
INSTITUTION Stanford Univ., Calif. Committee on Linguistics.
PUB DATE Nov 69
NOTE 39p.

EDRS PRICE MF-$0.76 HC-$1.95 PLUS POSTAGE
DESCRIPTORS Deep Structure; *Determiners (Languages); Form Classes (Languages); Grammar; Language Universals; *Nominals; Numbers; Phrase Structure; *Pronouns; *Sentence Structure; Structural Analysis; Suprasegmentals; *Surface Structure; Syntax

ABSTRACT

This paper argues that the hypothesis that all languages have a definitization process is empirically refutable, and that use of the terminology "definite" and "indefinite" is justified in that it reflects intuitions of grammarians and linguists. The following statements are tested against evidence from samples of different languages: (1) all noun phrases in a discourse may undergo definitization and turn out definite or indefinite; (2) all pronouns may similarly be marked for definitiveness; (3) at least some relative construction heads are definite; (4) numerated nouns, unless they have undergone the definitization, are indefinite. Following this, there is a discussion of the notion of "article" and how it may be described meaningfully. Appendices include tables of Indonesian articles in relation to pronouns, and of cross-linguistic distribution of some definitiveness markers. (Author/AM)

ED105717

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

DETERMINATION*

Edith A. Moravcsik

FL00 6293

*I would like to thank Miss Rebecca Agheyisi for her help in setting up the bibliography; Mrs. Dal Dresser for her editorial work; Professor Charles A. Ferguson and Dr. Russell Ultan for their comments, both substantive and stylistic; and, most of all, Dr. Alan Bell and Professor Joseph H. Greenberg for their patience and interest in discussing questions of determination with me.

INTRODUCTION

1.1 Let us hypothesize that all human languages have a definitization process, which may be tentatively and roughly characterized as a grammatical rule whose domain is the discourse and which obligatorily applies whenever two underlying simple predications contain noun phrases marked for identity of reference. The rule then copies the predication contained in one sentence into the noun phrase of the other. The noun phrase which thus acquires an embedded clause and which eventually gets ordered as subsequent to the other is a definite noun phrase. For example, from an underlying structure paraphrasable as

Something X is a girl. Something X is reading.
Something X is laughing.

we would get

A girl X is reading. A girl X is laughing.

and eventually

A girl X is reading. A girl X a girl X is reading is laughing.

and

A girl is reading. The girl who is reading is laughing.

The embedded sentence or its original is optionally subjected to a general deletion rule which deletes one of two identical constituents.

Two things need to be shown next. One is that the claim made above, according to which all languages have a definitization rule, is empirically refutable. Second, it should be shown that, given the meaning of the terms "definite" and "indefinite" as they are generally assumed by linguists, using these terms in the sense characterized above does not imply some arbitrary and unnecessary terminological innovation.

In terms of what was said above, a noun is definite if it is the same as one mentioned before; and definiteness marking is an optional surface manifestation of "sameness". Thus, we may construct a language which provides for no way of making clear whether one speaks about the same object or about another one. In other words, to make the claim that all languages have definitization is to exclude from the set of possible human languages a system where there is no way of saying two things about the same individual.[1]

---

[1] For an actual example of such an invented language, see Der schlaue Fuchs by István Fekete (Budapest, 1968). This novel consists of an animal story in which animals recognize individuals of their own race only and use unindividualized class terms for any member of another race, e.g. Gaga ist hier gewesen means 'Some member of the class Gaga (= ducks) was here'; it is irrelevant whether or not the speaker has seen that particular duck before. Examples are on pp. 52, 54, 56, 57, and passim.

As for justification of the terminology, it seems that to use the term "definite" for "something mentioned before" reflects some of the intuitions of grammar writers. In listing conditions for use of the definite article, the statement according to which "second mentionings" are definite has a prominent place in most grammars.

The hypothesized definitization rule will not be explored with respect to all its properties and implications. As it now stands, it needs justification with respect to the assumed principles it rests on--among others, that the domain of grammars is the discourse, that the order of constituents is predicted rather than underlying, and that referential identity is necessary and sufficient for a proper account of definiteness. No attempt to discuss these claims is being made here, however.[2] It is simply assumed that such an account is theoretically justifiable; and this paper is concerned only about factual adequacy.

Before considering various testable implications of the proposed rule, one obvious way will be pointed out in which this account appears to fail grossly. Since the rule operates on two noun phrases, the definiteness involved in discourses such as "Feed the cat!" remain unexplained--cases, that is, where the "indefinite antecedent" is situationally, rather than linguistically, given.[3] Such discourses are considered here ungrammatical; we assume that no grammatical discourse can begin with a definite (common) noun. The factual basis of this statement is that any such truncated or semidiscourse which starts with a definite common noun can be placed in a wider discursive context by any competent speaker of the language, in which case an overt indefinite antecedent would be supplied. The same kind of assumption is also indispensable if one argues for the predictability of pronominal agreement. The question, under what conditions and in what ways discourse may be truncated is a problem which, at the present stage of linguistic theorizing, must be separated from that of giving a linguistic account of definitization.

From the conjunction of our hypothesized definitization rule and from some other assumptions specified below, four particular claims will be derived for more detailed investigation:

1. Since there is no object that must not be referred to more than once in a discourse, it follows that all noun phrases in a discourse may undergo definitization and thus turn out to be either definite or indefinite.

---

[2] For arguments concerning the necessity and feasibility of discourse grammars and of axiomatic unorderedness, see Sanders 1967. On questions of what constitutes the kind of identity relevant for definitization and reduction rules, and how this identity should be marked, see Chomsky 1965, pp. 145-46; Sanders 1967, sec. 4.7; Lakoff 1968a, pp. 53ff. and 1968b, pp. 15ff. and 45ff.; Karttunen 1968; and Sampson 1969.

[3] I am indebted to Professor Joseph Greenberg and to Mr. Alan Bell for calling my attention to this point.

2. If all noun phrases are marked for definiteness and if pronouns are reduced noun phrases containing only the features which all noun phrases are marked for, the corresponding empirical hypothesis is that all pronouns are marked for definiteness.

3. By simply rephrasing the definitization rule, which contained relativization as a process responsible for definitization, we arrive at the hypothesis that (at least some) relative construction heads are definite and that the noun phrase in the constituent sentence is indefinite.

4. Accepting the suggestion that numerals are derived from "simple, nonnumerated sentence coordination" (Sanders 1967, sec. 4.5)--that is to say, "three pens" is derived from the conjunction of three first mentionings of three different objects, each named "pen"--it would follow that numerated nouns, unless they have undergone definitization, are indefinite.

In the next four sections of this paper, these statements will be tested against evidence coming from a sample of languages.

NOUNS

2.1 In this section we will first survey the various ways in which languages mark "second mentionings" of nouns. Subsequently, we will analyze the general distribution of these markers and will discuss cases where they appear to be distributed against our prediction.

Three types of formal devices will be distinguished:

1. Order
2. Segmental markers
3. Nonsegmental phonological features

1. Order - Whenever order correlates with definiteness, the definite element tends to be placed at the beginning of the sentence. In NORTHERN PEKINGESE, definite subjects stand before the verb. In LATVIAN, where order is the only obligatory definiteness marker for simple noun phrases, first position in the sentence signals definiteness. Definiteness and word order seem to be connected also in RUSSIAN, in FINNISH, in HUNGARIAN, and in SYRIAN ARABIC; the clue to all this may be laws of topicalization. What appears clear is that (surface) subjects and topics (in the sense of "assumed information") tend to be definite; that is to say, if topic or subject is restricted with respect to definiteness, the restriction is that they must not be indefinite (e.g. in TAGALOG and LEBANESE ARABIC, respectively). A third factor, also in some mysterious relationship with definiteness and with topicalization and ordering, is negation. In some languages, the indefinite marker does not occur (EGYPTIAN ARABIC, SAHIDIC COPTIC) or is fused (GERMAN) with the negative marker; in LUGANDA, the preprefix, which is in some respects a definiteness marker. does not appear after a negated verb.

Trubetzkoy's claim (1939, p. 137), according to which suspension of the definiteness contrast for the subject but retention for the predicate noun occurs in no human language, thus seems to be well-founded; however, it is also interesting to notice that the category which is apparently most free to exhibit contrasts in definiteness is the (surface) object. Even if a language has other obligatory definiteness marking, the definite object seems, in a number of languages, privileged to have an additional pronominal marker. Furthermore, there are languages which have no obligatory definiteness marker for any noun other than objects. This is said to be true for TURKISH, MODERN PERSIAN, some GREEK dialects of Asia Minor and, possibly, for TANNA and MALOESE. It also holds for one noun class in BENGALI. In CHIRICAHUA, obligatorily expressed definiteness contrast exists for the third person object pronoun only. In KABARDIAN, contrast exists only for the subject of the intransitive verb and for the object. Considering all this, the following statement seems to hold true: for any particular language, definiteness contrast obligatorily expressed for any noun phrase implies such contrast in the object noun phrase.

2. Segmental markers - Evidence appears to allow for establishing three definiteness markers as options for speakers of all languages included in the sample, and possibly for all other languages as well. One is the restrictive relative clause--a marker that is directly predicted by the definitization rule. Although grammars of several languages did not deal with this question, no statement has been found anywhere claiming that nothing comparable to a relative clause exists in a given language; therefore, let us assume that the restrictive relative clause is a universal construction. Second, existence of "sameness and/or otherness attributes" has been observed for a number of languages. In these, adjectives meaning "same" do not usually co-occur with an indefinite article, but may with a restrictive relative clause; they often co-occur with or bear partial resemblance to some other definiteness marker. In particular, such adjectives have been found to morphologically "contain" either a demonstrative or a definite article in the following languages: CLASSICAL ARABIC, COPTIC, FULANI, FUTUNA, HUNGARIAN, NORTHERN PEKINGESE, as well as most, if not all, INDOEUROPEAN languages. Third, a demonstrative can be used in a "referential" rather than locative sense, with the meaning 'aforementioned', in a number of languages, e.g. ACHOLI, CHILUBA, CHITIMACHA, CORNISH, FULANI, HAUSA, HUNGARIAN, ICELANDIC, LUGANDA, MAASAI, NORTHERN PEKINGESE, PORTUGUESE, ROMANIAN, SANSKRIT, TLINGIT, TONKAWA; WOLOF. These anaphoric demonstratives must never occur with an indefinite article, but they may co-occur with a restrictive relative clause; in the majority of cases they are somewhat similar to a definite article.

Besides these optional markers, a number of obligatory markers have been observed with a restricted cross-language distribution. In particular, the demonstrative is apparently not the only pronominal element that occurs optionally or obligatorily with definite nouns. In NORTHERN PEKINGESE, and also in ICELANDIC, a third person pronoun co-occurring with a noun may mark its definiteness (although it is not clear whether the ICELANDIC pronoun is a redundant marker standing with proper names only or may occur with common nouns too);

in FRENCH, however, the object form of the contemporary third person singular pronoun is an obligatory definiteness marker. There is a particular tendency for third person pronouns to co-occur obligatorily with definite nouns in the object case as definiteness markers. This is true in MACEDONIAN, NORTHERN PEKINGESE, and LEBANESE ARABIC; it has been suggested as the best way to account for object incorporation in the verb in HUNGARIAN and, possibly, in ALBANIAN and KWAKIUTL. Besides the verb, the adjective, too, may contain definiteness markers, as in LATVIAN, ICELANDIC, GERMAN, BULGARIAN, LUGANDA, and ARABIC. It is interesting to note, however, that markers on the verb or adjective are never the sole markers of definiteness, in any given context (except perhaps in LATVIAN); their presence generally implies that of some other definiteness marker.

Furthermore, a particle generally called the "definite article" has been found to mark second mentionings in terms of an obligatory rule. Because of the widespread phonological similarity of the respective elements, the domain of the statement "definite articles mark definiteness" is partially covered by our other two previous statements, according to which demonstrative pronouns (optionally) and third person pronouns (obligatorily) mark definiteness. Nevertheless, some evidence, to be discussed in the final section, shows that even if all definite articles did indeed look like either of these two pronouns, we could still not take them to be quite the same thing.

Finally, an "indefinite article", also to be discussed later, marks indefiniteness in certain languages.

3. Nonsegmental phonological features - Tone has been found to mark the definiteness contrast obligatorily in BAMBARA, in CHILUBA, and in LOMA. As for stress, all segmental obligatory definiteness and indefiniteness markers (such as pronouns, definite and indefinite articles) are unstressed, unless they have contrastive stress.

2.2 Having surveyed various second- and first-mention markers, we will now discuss cases in which these markers do not occur as predicted or their distribution is not accounted for by our definitization rule. As a guideline in our discussion, consider ENGLISH. In this language we would expect the obligatory indefiniteness marker--the indefinite article--always to appear with first mentionings of nouns, and the obligatory definiteness marker always to co-occur in additional mentionings. Now consider the following discourses:

1. Do you see that book? Take that book with you.
2. My book is on the table. Why don't you take my book?
3. John is a good friend of mine. John is an engineer.
4. A sea lion is a big animal. A sea lion does not like hot water.
5. Sir, what time is it? Thank you, Sir.
6. She is an anthropologist. She never wanted to become an anthropologist.
7. She entered with joy and departed with joy.
8. John rented a typewriter. He hates that typewriter.[4]

[4]For most speakers, some of these sentences require pronominalization.

In each case, the first sentence may conceivably start a discourse; the underlined repeated noun phrases do not refer to clearly different objects in 4, 6, 7, and 8; and they do refer to the same object in the rest of the sentences. Still, they either have no articles or there is no switch from one article to the other during the discourse. The definitization rule appears to apply in none of these discourses.

2.2.1 As for noun phrases containing a demonstrative, there is much evidence to prove that such constructions are treated as definite in all languages. We have already seen above that some demonstratives--those devoid of reference to spatial relations--are simply optional definiteness markers. Now we will see that in fact all demonstratives render a noun phrase definite; it must be explained why this is so.

We infer the definiteness of demonstrative constructions from the following facts. In NORTHERN PEKINGESE, such constructions have the same order privileges as noun phrases mentioned for the second time. In ARABIC, the relative pronoun or the adjective, both indicating the definiteness status of the head, bear the definiteness marker when the head is a demonstrative construction. In HUNGARIAN, the verb ending, symptomatic of the definiteness of its object, has definite endings if the object is a demonstrative construction. In ICELANDIC and GERMAN, weak adjectives occur in such constructions--i.e. adjectives with the same endings that occur in constructions containing the definite article. Furthermore, many definite articles are historically related to demonstratives of the same or of a related language and are partially similar or identical with them.

Thus, the question arises: given one single mechanism by which noun phrases may become definite, how do we account for the definiteness of demonstrative constructions? An interesting fact about such noun phrases is that, apart from some dubious cases, they may not be heads of restrictive relative clauses. This fact, along with numerous others referred to above showing the definiteness of these constructions, is accounted for if we assume that a restrictive relative clause is already in the noun phrase. That is to say, we derive "this apple" from "the apple which is here". It is presumably true for all languages that these two constructions are both possible and synonymous with each other --a fact which this treatment accounts for also.

2.2.2 In the second sentence above, the definitization rule does not appear to hold for possessed noun phrases. Surveying data from various languages concerning possessive constructions, the following generalities emerge. The possessed noun is usually treated as a definite noun. In LATVIAN, the long, i.e. definite adjective form is used with it. In HUNGARIAN, the verb having a possessed noun as its object takes definite endings. As for articles, if the possessor is a pronoun there is either no article (ARABIC, CELTIC, ICELANDIC--when the possessive pronoun stands before the noun, BULGARIAN, FAROESE, LUGANDA, TUNICA, WELSH) or the definite article or some optional definiteness marker is used (ALBANIAN, COPTIC, ATHAPASCAN, LOMA, PORTUGUESE, PANGKUMU--in the four latter languages this is the case for

alienable but not for inalienable possessions, ICELANDIC--if the possessive pronoun follows the noun, BENGALI, BIERIAN, DANISH, FIJIAN, HUNGARIAN, MACEDONIAN, MAE, OLD PERSIAN, SOMALI, TIGRE. For some languages, therefore, the definite article does not alternate with the indefinite one in such contexts (rather, if the possessed item is to be indefinite, some other construction is used); however, in other languages there is such a simple alternation of articles, as in BAMBARA. HUNGARIAN provides an interesting case: although both definite and indefinite articles may co-occur with a possessed noun, the verb shows the definite ending if the possessed noun is an object, even if it is marked indefinite by the article and is so understood.

Since it is presumably true for all languages that "my book" is synonymous and freely interchangeable with "the book which is mine", it is assumed that possessive constructions, too, are derived by embedding and that this process accounts for the definiteness of the possessed item. It is worth mentioning that in ATESO use of the relative construction is the only way for a noun and a possessive pronoun to constitute a noun phrase; apparently no language uses only the reduced construction. Since our rule predicts definiteness for the head of a relative clause, there is no way to account for the fact that in some languages the possessed noun can be indefinite; this question must be left open (compare sec. 4). As a final comment on possessives, let it be pointed out that just as it is probably true for all languages that in some contexts a definite article and a demonstrative vary freely, so too may possessive pronouns, presumably in all languages ("I looked at the house. The/its door was open.").

2.2.3 Proper names are generally treated as definite noun phrases. In ARABIC and in ICELANDIC, adjectives show definiteness if they modify a proper name; in HUNGARIAN, the verb does likewise if the proper name is an object. In most languages, proper names do not have segmental definiteness markers (CELTIC, HIDATSA); in others, some have definite articles, some do not (RUMANIAN, BULGARIAN, MODERN GREEK, SWISS GERMAN, ARABIC, COPTIC). In ANCIENT GREEK, the proper name with the definite article has occasionally some scornful connotation. CLASSICAL ARABIC was the only language found where proper names may have an indefinite article without losing their unique reference. But whatever articles proper names take, it is characteristic that they do not change articles, as sentence 3 on p. 68 indicates for ENGLISH. Since the definiteness of proper names is therefore beyond the scope of the definitization rule--they neither co-occur with nor imply any restrictive relative clauses--we conclude that proper names are inherently definite.

In several languages of the Pacific Islands, a particle which grammars call "personal article" obligatorily co-occurs with proper names. If, however, we characterize articles as elements obligatory for a particular language that mark one or the other member of a definiteness contrast, these proper name markers do not qualify as articles since they are not assigned by a definitization process. The only feature that distinguishes them from any arbitrary particle is that almost all of them have some other function in the same or in a related language --invariably that of a demonstrative or personal pronoun or of a "real" definite article. See Appendix I for a pronominal function chart of Indonesian articles.

2.2.4 In sentence 4 on p. 68, "a sea lion" stayed unchanged when mentioned for the second time; the question is, why? We might hypothesize that this noun is just like proper names: it is inherently definite; or it is inherently indefinite. However, this cannot be true since phrases such as "sea lion" do undergo definitization in other sentences. Remembering that the condition under which the definitization rule was supposed to operate was identity of reference, we might argue that the two mentions of "a sea lion" do not have the same reference, and if not but they do refer to something, they must refer to different things. To see whether this is true we might observe the syntactic behavior of such nouns with respect to a syntactic rule which operates in terms of nonidentity of reference. Such a rule would be pluralization (on pluralization as a syntactic process, see Sanders 1967, sec. 4.10). From "A man is in the room. A(nother) man is in the room", where the two "man"-s have different references, we get "(Two) men are in the room." This process seems inapplicable to discourses such as "A sea lion does not like hot water. A sea lion does not like hot water."; these two sentences do not reduce to "Two sea lions do not like hot water", unless by "a sea lion" we mean one particular one, thus departing from the natural interpretation of discourse 4 on p. 68. Since, therefore, such nouns apparently have neither the same reference nor nonidentical references, we conclude that they have no reference at all, thus making questions concerning "sameness" and "otherness" inapplicable (this conclusion about generic noun phrases was reached by Baker 1966, pp. 19ff.). Accordingly, the hypothesis that all nouns are either definite or indefinite has to be modified. Nouns are now classified as generic and specific; the latter are either definite or indefinite, in terms of inherent marking or prediction.

Again, it might be in order to consider that it is not an a priori necessity for all languages to have both generic and specific nouns. One might conceive of a language in which nouns are used in a specific sense only; on the other hand, a language with generic nouns only would consist solely of definitions and aphoristic sentences.

According to this feature system, it would not be surprising to find that, for marking generic nouns, languages employ a system which bears no relationship to that utilized for marking definite and indefinite nouns. In reality, this is not so at all. No language with articles has been found where a generic noun could not be marked by means that are also employed in definiteness marking; therefore, our system is deficient in failing to exclude the possibility of complete independence of the two sets of markings.

Statements in grammars about marking generic nouns are few and incomplete. In languages that have no article, "all" and "every" or "any" might co-occur with nouns that are meant to be generic; verb aspect may also be a marker (in ENGLISH, in some contexts, the progressive does not go with generic nouns). In languages with articles, these or their absence may also mark generic and other nouns. In ENGLISH and HUNGARIAN, all possibilities with respect to number and articles--that is, singular and plural indefinite and definite noun phrases--may be generic, although certain genericity markers seem to be restricted to nouns with few semantic features ("Man will soon conquer the moon"

as opposed to "*Girl will soon conquer the moon"). But whatever their marker (BAMBARA, singular definite or indefinite; COPTIC, singular definite or indefinite, plural definite; ANCIENT GREEK, singular definite; BULGARIAN, ICELANDIC, no article; LUGANDA, MAASAI, no preprefix; SYRIAN ARABIC, singular definite) and whatever construction they may happen to constitute (generic noun phrases may contain a restrictive relative clause and may even be possessive constructions--"My son shall never marry a French girl."--a demonstrative construction is the only one which cannot be used generically), one generality holds for all: no generic noun changes articles, i.e. they do not undergo definitization. In fact, we might posit a more general statement: generic nouns undergo no syntactic process that requires identity or nonidentity of reference; the one apparent exception to this is that generic nouns are pronominalizable ("A sea lion is a big animal. It does not like hot water."). Of course, the existence of generic pronouns could have been guessed when we established that nouns are either specific or generic: features which all nouns must be marked for are presumably also pronominal. Since in general pronominalization works in terms of identity of reference, and according to our system generic nouns have no referential marking, we must assume that pronominalization of generic nouns works in terms of identity of noun phrases.

As we have just observed, pluralizability (or numeratability) and definitizability are related; in particular, the former implies the latter but not vice versa (see "the water", "the patience that..." but not "*two waters", "*patiences"). This implicational relationship would remain unexplained if we did not assume that both numeration and definitization presuppose referential marking. Assuming, however, that numeration works in terms of nonidentity and definitization in terms of identity of reference, the correlation is explained. That the relationship is a hierarchical rather than a mutual implication is because definitization requires referential marking only, numeration also requires the noun to be marked as count.

2.2.5 Considering example 5 on page 68, which contains vocatives invariable under repeated mentioning, it is important to realize that, as in the other examples, we face a universal phenomenon: vocatives do not undergo definitization in any language investigated. Most vocatives take no article at all or take the definite article. There is no article in EGYPTIAN ARABIC, HOWA, BULGARIAN, PORTUGUESE, MAASAI, ICELANDIC, HUNGARIAN; in ICELANDIC, adjective agreement shows that the noun is definite. The only language in the sample where vocatives take the definite article is COPTIC. Since vocatives invariably pronominalize in the second person (or, rather, in the person of the hearer), but not in the third person (unless the third person includes a "hearer's person"), we conclude that only "third person nouns" undergo definitization and "second person nouns" are definite because the second person pronoun is inherently definite.

2.2.6 Just as example 6 shows for ENGLISH, predicate nouns do not undergo definitization in any language of the sample. Predicate nominals in general seem to be indefinitely marked, if anything (many predicate adjectives in NORTHERN PEKINGESE have the indefinite marker), but in some languages they must not take the indefinite article (HUNGARIAN). In BULGARIAN and

ICELANDIC, the predicate adjective always has the indefinite form (i.e. the "short" or "strong" form, respectively). In ARABIC, the predicate adjective is usually without article; in LUGANDA, it takes no preprefix; in LOMA, it never occurs with a definite suffix. Predicate adjectives are often similarly restricted with respect to the marking of other categories, such as gender and number; for predicate nouns, too, a cluster of properties all show deviation from the behavior of nonpredicate noun phrases. While predicate nouns are not definitized, neither are they pluralized (further confirmation of the statement "Numeratability implies definitizability"); furthermore, they do not commonly pronominalize (or rather, if they do, the resulting pronoun has a restricted distribution as it, too, must be a predicate; see "She is an anthropologist. She never wanted to become one.", not "*She is an anthropologist. He/she/it is a nice girl." where he/she/it refers to "an anthropologist"); and they cannot take nonrestrictive relative clauses. These two latter properties distinguish them from generic nouns. Since predicate nouns behave differently from both specific and generic nouns, we conclude, accepting Bach's recent proposal (1968), that predicate nouns are not noun phrases at all. This conclusion, of course, leaves unexplained the fact that predicate nominals do acquire indefiniteness and definiteness markers, as opposed to, say, verbs; it also leaves the status of predicate pronouns unclarified. These questions must be left open here.

From data relating to generic nouns we have inferred that the domains of pronominalization and definitization are not coextensive; generic nouns can be pronominalized but not definitized. Vocatives, too, can be pronominalized but are not subject to the definitization rule. With respect to predicate nouns, we have just seen that they can be neither pronominalized nor definitized. Thus, we tentatively assume that, whereas pronominalizability does not imply definitizability, the latter does imply the former for any given noun phrase (for some exceptions to this in ENGLISH, see P. Wolfe 1968). On the other hand, we have also seen that numeratability implies definitizability but not vice versa. Thus, the distributional properties of these three processes relative to each other can be summarized as follows: for any given noun phrase, numeratability implies definitizability, and definitizability implies pronominalizability. In other words, noun phrases that can be pluralized but not definitized and those that can be pluralized and definitized but not pronominalized are excluded as possible constructions.

2.2.7 These relationships may help to explain why nouns occurring in adverbial phrases and in compounds do not become definitized, as the last two sentences on page 68 show for ENGLISH. Probably all languages with articles have certain stereotyped adverbial phrases which may or may not contain an article; but there is no change of article during a discourse. Such constructions are in EGYPTIAN ARABIC those glossed as "this morning" and "from house to house"; in BULGARIAN, "per day", "all day", "for the first time"; in RUMANIAN, adverbs such as "with speed",or "to sit on a chair"; in YIDDISH, "to school"; in FRENCH, "last night", etc. In RUMANIAN, no article is ever used with prepositions unless the noun phrase also contains an adjective. In LUGANDA and MAASAI, the preprefix is generally not used in adverbs. In COPTIC, certain adverbial phrases such as "wisely", "fearfully", "with

knowledge", contain an indefinite article and others, such as "in the field", "before men", "without loot", have definite articles. The same is generally true for nouns included in compounds; they usually lack definiteness markings (in MAASAI, they lack prefixes), or else these are "petrified" as in ENGLISH "Jack-in-the-box" or "son-of-a-bitch". But given the fact that neither can such nouns be pronominalized, this becomes a property derivable from our proposed implicational hierarchy. In terms of that hierarchy, it is predicted that neither do these nouns undergo pluralization; this can be easily proved. The same reasoning applies to inner objects ("He died a painful death").[5]

2.2.8 In conclusion: we first surveyed various optional and obligatory definiteness markers and concluded that all languages can optionally mark definite nouns in terms of a restrictive relative clause (which is obligatory in deep structure), or by an anaphoric demonstrative or third person pronoun, or by "sameness-otherness attributes". Other markers turned out to be obligatory and nonuniversal. In order to test the hypothesis "All nouns are either definite or indefinite", we surveyed the distribution of these various definiteness markers. We concluded that demonstrative and possessive constructions should be derived by embedding. We also concluded that the following feature system for marking nouns should replace the definite-indefinite dichotomy: some nouns are marked as inherently definite (and, redundantly, specific); others may be used as either generic (i.e. without referential marking) or indefinite (which implies specificity). Indefinite nouns have referential marking; since they are pronominalizable in the third person, they must undergo definitization if they refer to the same thing more than once. A proposed implicational statement claims that whereas referential marking is not a prerequisite for pronominalization--even generic noun phrases and "predicate nouns" can be pronominalized --only a subset of pronominalizable things, i.e. those which have reference, can undergo definitization; and (though all nouns with reference are quantifiable) only that subset which is countable can be numerated.

PRONOUNS

3.1 From a semantic point of view, pronouns may be roughly grouped as generic, definite, and indefinite. In "People are cruel. They do not consider each other's interests", "they" is generic. "Somebody" and "who"-type pronouns do not refer to someone familiar to the speaker; "he", "his", "himself", "this (one)" refer to persons and objects already known; so do "I", "we", and "you", although these are not substitutes for any noun. Finally, there is the relative pronoun; whereas people's intuitions would probably be in agreement on the above classification, they are likely to differ on how to classify "who" or "which".

---

[5]Mr. Alan Bell called my attention to inner objects.

Generic pronouns will be discussed no further in what follows, but generally it seems that any personal pronoun, except perhaps "I", can be used generically.[6] Nor will we discuss predicate pronouns that were mentioned in the preceding sections. We will consider only "definite" and "indefinite" pronouns to see if there is any syntactic evidence for such definiteness marking with respect to pronouns. Evidence will be sought by comparing the distribution of pronouns with that of definite and indefinite noun phrases and by considering the phonological shape of pronouns.

Distributional evidence might point to any of the following conclusions:

1. The definiteness feature is not a pronominal category.
2. The definiteness feature is a pronominal category and (some) pronouns are like the majority of nouns: pairs of them contrast only in terms of the definiteness feature.
3. The definiteness feature is a pronominal category and (some) pronouns contrast in terms of more than one feature, including definiteness; language-specific redundancy rules predict the value of definiteness.
4. The definiteness feature is a pronominal category and (some) pronouns contrast in terms of more than one feature, including definiteness; the redundancy rules in terms of which the value of the definiteness feature is predictable are the same for significant groups of languages.

If we find that pronouns pattern at random, in comparison with the distribution of definite and indefinite noun phrases, we will say the 1. is true. If we find pairs of pronouns that are distributed like pairs of nouns with the same meaning but differ in definiteness, this will be taken as evidence for 2. If it is found that pronoun X is distributed like a definite noun in one language and like an indefinite noun in other languages, we will conclude at 3. If the distribution of pronoun X is the same as that of either definite or indefinite nouns, for any group or for all languages, we will tentatively accept 4.

3.2 Personal pronouns often "contain"--or obligatorily co-occur with--some definiteness marker which is either a demonstrative or a definite article. The first person singular pronoun in EROMANGA, the first and third person singular pronoun in EFATE, FUTUNA, TANGOAN, and MALAY, the third person singular in MALAGASY, and the third person plural in AMHARIC, contain a demonstrative. (See also Appendix I.) The third person pronouns in the singular are partially identical with the definite article in MALAGASY, FRENCH, PORTUGUESE, BIERIAN. In FIJIAN, the personal article which co-occurs otherwise only with proper names also occurs with all personal pronouns.

---

[6] As far as generically interpreted possessive pronouns are concerned, in CHIPEWYAN and in AZTEC special possessive pronouns ("someone's") occur with these; in AZTEC, plural first person pronouns, too, may be used in such a generic sense, just as in CHIQUITA. In CHOROTI, "his", which also means "animal's", is used with generically-employed animal possession. In ALGONQUIAN also the third person pronoun is used; in CHIRICAHUA is is the third person pronoun, jointly with some other element.

In WELSH, all personal pronouns co-occur with the definite article. The phonological shape of the third person pronoun is identical with that of a demonstrative (which is usually far-pointing) in SAMOAN, FIJIAN, BULGARIAN, ICELANDIC, ATHAPASCAN, ALEUT, ALBANIAN, FULANI, LOMA, and NORTHERN PEKINGESE. (No example has been found, however, of a demonstrative functioning as a non-third-person pronoun, e.g. "this" for "I".) Adding the fact that in NORTHERN PEKINGESE all personal pronouns are ordered as definite nouns, we have evidence for the statement: All personal pronouns are definite. Some other facts, however, seem explainable only if we assume that some personal pronouns are indefinite. In HUNGARIAN when the verb has a third person pronoun for an object, it has the same suffix series as with definite objects; but it has "indefinite endings" when the object pronoun is in the first or second person. Furthermore, some evidence also shows that pronouns of the same person may constitute contrasting pairs with respect to definiteness. In LOMA such a distinction is claimed for first and second person pronouns: the definite forms occur as pointers and heads of relative clauses, the indefinite ones elsewhere. In BUG and HOWA, the first person singular pronoun may or may not be prefixed with a demonstrative; when it is so prefixed it is more "emphatic". In striking contrast to the scarcity of evidence for such minimal pairs for non-third-person pronouns, minimal pairs exist for the third person singular in all languages of the sample: all languages have an equivalent to "someone" next to another form corresponding to "he/she/it".

3.3 Possessive pronouns (used nominally) have nowhere been found to appear in an environment characteristic of indefinite noun phrases; they obligatorily co-occur with definiteness markers in, among others, PORTUGUESE, in HUNGARIAN, and in some SCANDINAVIAN languages.

3.4 We have little data on reflexives. Whenever they have some marker relatable to definiteness, it always indicates the positive value of this feature. In LEBANESE ARABIC when a reinforcing function co-occurs with another noun phrase, it co-occurs with a definite article.

3.5 At least a binary proximity contrast for demonstratives has been found in all languages of the sample; contrasting with possessive pronouns, the same form may occur independently, i.e. as a noun or accompanying a noun. Evidence for the definiteness of all demonstratives[7] is the fact that they are derivable from "second mentionings" and that they often co-occur with a definiteness marker (in EFATE, SAMOAN, SANTO, TANNA, MACEDONIAN, MAASAI, LUGANDA, LOMA, LYELE, BENGALI, ATHAPASCAN, EGYPTIAN ARABIC, IRISH, ANCIENT GREEK, HUNGARIAN) or with a definite verb paradigm (in HUNGARIAN). Furthermore, the definite article itself is, in most cases, originally or even synchronically identical with a demonstrative.

---

[7] Although "such a"-type pronouns might simply turn out to be indefinite demonstratives.

3.6 As was mentioned in connection with personal pronouns, all languages of the sample have an indefinite (third person) pronoun. It is worth pointing out, however, that in no language does this third person indefinite pronoun show any phonological similarity--a common stem, say--with the definite third person pronoun, as similar minimal pairs of nouns do. Another difference is that the definite third person pronoun has either the same associated categories (i.e. number and gender) as its indefinite counterpart or more, but never fewer.

3.7 The indefinite pronoun has been found to be identical or similar, in most cases, to some question pronoun. If they are partially similar it is the indefinite pronoun whose shape includes morphologically that of the question pronoun. Languages with this similarity between the two pronouns are the following: ARABIC, EFATE, CHITIMACHA, MALAGASY, ALBANIAN, BENGALI, ICELANDIC, MACEDONIAN, FUTUNA, SAMOAN, BAKI, MALOESE, BIERIAN, MALEKULA, RUMENIAN, DAKOTA, TONKAWA, BULGARIAN, CHIRICAHUA, HOPI. CHIPEWYAN, IRISH, NORTHERN PEKINGESE, COPTIC, PANGKUMU, FULANI. Although question pronouns usually do not occur with an indefinite marker (although they may optionally in SAHIDIC COPTIC), some definiteness marker is morphologically contained in question pronouns of the following languages: TANNA, EFATE, FULANI, ARABIC, ALGONQUIAN, BENGALI, BAKI, DAKOTA, AMHARIC, LOMA, MALEKULA, FUTUNA, SAMOAN, MALAGASY, MALOESE, BIERIAN, MALAY, FIJIAN, ANCIENT GREEK. NORTHERN PEKINGESE word order classifies question pronouns with indefinite nouns; so do HUNGARIAN verb endings, excepting "which?"--it is definite. On the other hand, in LEBANESE ARABIC question pronouns can be subjects, although there is a rule that indefinite noun phrases must never be subjects.

3.8 We conclude that nongeneric pronouns prove to be either definite or indefinite; that there are minimal pairs for the third person singular in all languages examined, and for other persons and perhaps even for demonstrative and question pronouns in some languages; and that the redundancy rule, which predicts definiteness for non-third-person, for possessive, and for reflexive pronouns, and for demonstratives, and indefiniteness for question pronouns, has some interlingual generality.

3.9 Things are not nearly so clear with respect to the relative pronoun. Given our pronoun-types and their definiteness markings as specified above, the following types of relative pronouns may be distinguished:

1. The relative clause is introduced by an indeclinable word; this type of clause, in some languages, also contains a "resumptive pronoun" referring to the antecedent. This pattern (but not necessarily only this) is illustrated by the following languages: ACHOLI, ALBANIAN, EGYPTIAN ARABIC, OLD ENGLISH, HEBREW, SAMNANI.

2. The relative pronoun is a unit which morphologically contains a demonstrative-definite article and a question/indefinite pronoun, e.g. ITALIAN, SPANISH, HUNGARIAN, BULGARIAN, ALBANIAN, EARLY MODERN ENGLISH, PORTUGUESE. In HUNGARIAN, the verb agrees with it as with an indefinite noun phrase.

3. The relative pronoun is identical with a question (/indefinite) pronoun, e.g. BENGALI, BULGARIAN, CHIPEWYAN, FULANI, MACEDONIAN, ALBANIAN, CLASSICAL ARABIC, MAIDU, TLINGIT, WELSH, PORTUGUESE, RUMENIAN. In MACEDONIAN, the definiteness-marker pronoun co-occurs with the relative pronoun when it is an object.[8]

4. The relative pronoun is identical in shape or closely similar to a definiteness marker

a. Marker a demonstrative: ALBANIAN, ALEUT, CLASSICAL and EGYPTIAN ARABIC, ARAMAIC, ATESO, ATHAPASCAN, BENGALI, BERBER, BIERIAN, DAKOTA, EFATE, FIJIAN, FULANI, MIDDLE IRISH, JAVANESE, MALAGASY, MALAY, MALOESE, PANGKUMU, RUKUMBU, SAMOAN, TANGOAN, TANNA, TSHIMSHIAN

b. Marker a third person pronoun: CHILUBA, FIJIAN, FULANI, FUTUNA, LOMA (where every personal pronoun has a relative form), PANGKUMU, OLD PERSIAN, RUKUMBU, WEASISI

c. Marker a definite article: SUDANIC ARABIC, HOMERIC GREEK, SAMOAN

d. In FIJIAN a possessive pronoun can also be a relative pronoun.

5. No pronoun-like or conjunction-like element is found in at least some relative clauses in EGYPTIAN ARABIC, BIERIAN, EFATE, FUTUNA, JAVANESE, MALAGASY, MALAY, MAIDU, MALOESE, PANGKUMU, NORTHERN PEKINGESE, RUKUMBU, SAMOAN, TANNA.

6. The relative pronoun is a preposition: EGYPTIAN ARABIC, BERBER.

7. The relative clause is introduced by some unidentified marker: BAMBARA, BIERIAN, CHIPEWYAN, ICELANDIC, LOMA, NAVAHO, NORTHERN PEKINGESE, WEASISI.

In some languages the relative pronoun has become the conjunction of other subordinated clauses. In ARABIC, it may introduce object clauses, as it may in ARAMAIC, FRENCH, and LATIN. In EFATE it may introduce causal and purpose clauses. In CELTIC "when" arose from a relative pronoun; in ATHAPASCAN "and" derives from a demonstrative.

Relative pronouns thus comprise morphologically both a definite and an indefinite element (1. and 2.), or only an indefinite element, or only a definite one. Like indefinite and question pronouns, they never have more categories associated with them than third person definite pronouns; they almost always have animate-inanimate distinction (not in ICELANDIC). It is also worth noting that, at least in one language, when grammatical and natural gender differ, the relative pronoun follows the grammatical gender, while the third person definite pronoun agrees with the grammatical gender or with the sex feature (see GERMAN "Das Mädchen das im Zimmer sitzt ist hübsch. Es/Sie ist meine Freundin.").

---

[8] Regarding the relationship of questions and relative clauses, LUGANDA, MAASAI, and CELTIC languages may provide some interesting evidence; it is said that in these languages questions always consist of a relative construction.

The fact that pronouns have been found to be definite or indefinite makes possible a tentative generalization about agreement. Nouns and verbs, adjectives, etc. clearly do not agree in terms of any conceivable semantic or syntactic feature that a noun may have; this fact should be captured by some generalization. It has been suggested that agreement might be described as consisting of transferring or copying a pronoun onto the agreeing element. For this to be at least somewhat plausible, agreement must never take place in terms of non-pronominal categories, i.e. those not present in pronouns. Since in some languages nouns and adjectives seem to agree in definiteness, it is important to recognize that definiteness is a pronominal category and therefore definiteness agreement does permit a description according to which the presence of a pronoun makes two terms agree.

MODIFICATION

4.1 According to our definitization rule, the identical noun phrase in the relative clause is indefinite. As we have just seen, the pronoun that replaces it is indeed often indefinite, and thus the rule explains these cases. However, in other instances the relative pronoun looks more like a definite pronoun or it contains morphologically both kinds; the definitization rule does not account for such pronouns. Second, as shown with respect to possessive constructions, the rule predicts that heads of relative constructions are definite; but no language has been found where the definiteness contrast, in actual utterances, would have been restricted in such a way. (For some hypotheses on the relationship between relativization and definitization, see Robbins 1961; Annear 1965; and Baker 1966.)

Both these facts might be explained by representing definitization as some recursive process. It appears that by adding a relative clause to a noun phrase, a subset of the set the head refers to becomes specified, i.e. definite with respect to the head-set. On the other hand, it can also be regarded as a genus of a new species, to be defined by some other relative clause. Since no actual rule with any motivation can be offered, we will simply conclude here that facts about relative clauses are more complicated than our present definitization rule can account for. In what follows, a set of simple observations on relatives will be presented in terms of a tentative taxonomy.

4.2 Structural properties of relative clauses depend on certain properties of the head, on the intrasentential function of the identical element in the subordinate clause, on the tense of the subordinate verb, and on the relationship between the two clauses. The following correlations are observed:

1. Head

a. Explicit-nonexplicit - in EGYPTIAN ARABIC, usually ma or nothing is used as a conjunction if the head is nonexplicit, ʔilli and ʔinni if it is explicit.

b. Definite-indefinite - in EGYPTIAN ARABIC and COPTIC, this governs the shape of the relative pronoun; in TANNA, if the head is indefinite there is no conjunction and none (or it is a demonstrative) if the head is definite;

In SPANISH and PORTUGUESE the subjunctive is used when the head is indefinite, in SYRIAN ARABIC it may be so used.

c. Vocative-nonvocative - in EGYPTIAN ARABIC there are different conjunctions depending on this feature.

2. Subordinate clause

a. Intrasententential function of identical element

(1) Subject or other - ACHOLI, ALBANIAN, SYRIAN ARABIC, SAHIDIC COPTIC, SAMNANI, SOMALI: resumptive pronoun used if nonsubject; LUGANDA: preprefix used if subject; SOMALI: subjunctive used if nonsubject

(2) Subject-object or other - OLD IRISH: relative pronoun used if subject-object, none otherwise except for possessive pronoun; FULANI: resumptive pronoun if other

(3) Object or other - EGYPTIAN ARABIC (and optionally in ENGLISH): no relative pronoun if object

b. Verb

(1) Pluperfect - no pronoun in EGYPTIAN ARABIC

(2) Future - special pronoun in LOMA

(3) Imperfect and present of habitude - BOHAIRIC COPTIC has indefinite relative pronoun

3. Intersentential relation

a. Restrictive-nonrestrictive - obligatory personal pronoun if nonrestrictive, optional otherwise in EGYPTIAN ARABIC

b. Indicative if nonrestrictive or if restrictive and relative pronoun is subject, subjunctive otherwise in SOMALI

c. Postposed relative clause if nonrestrictive, preposed otherwise in BAMBARA

d. No pause if restrictive, pause between clauses otherwise, presumably in all languages.

BAMBARA and BENGALI are the only languages found to have obligatorily preposed relative clauses. It is interesting to note another property that is unique to these two languages: it is the head that is substituted by a pronoun and the relative marker co-occurs with the noun ("...which N...it..."). Thus, there seems to be a universal tendency to reduce that noun that is in second place, regardless of the order of the two sentences--a fact easily accounted for in terms of a general second-mention deletion rule. The corresponding tentative universal is that in a relative clause, the second mentioning of the identical element always has fewer features than the first.[9] Notice also that in these two languages the deleted head is always substituted by a definite pronoun.

---

[9]Notice also ENGLISH "I recommended Gulliver's Travels to him, which book you have already read" but not "*I recommended the book to him which Gulliver's Travels you have already read."

Apart from BAMBARA and BENGALI, the following generalization can be made for all languages of the sample with respect to ordering of the two clauses: For any language, preposed relative clause implies postposed relative clause. This correlates, to some extent, with Greenberg's proposed universal about adjective order: "When the general rule is that the descriptive adjective follows, there may be a minority of adjectives which usually precede, but when the general rule is that the descriptive adjectives precede, there are no exceptions" (1963, p. 87).

Some other generalizations, apparently exceptionless, can be abstracted from the above classification:

Occurrence of a resumptive pronoun in the relative clause in cases where the identical element is subject or object of the clause implies its occurrence in cases where the relative pronoun has some other function.

Occurrence of a resumptive pronoun in a relative clause implies that the relative pronoun is invariable.

If there is an indicative-subjunctive contrast in relative clauses with the choice depending on the relationship of the two sentences, it is the restrictive clause that will have the subjunctive.[10]

For any language, an obligatorily definite identical noun phrase in the preposed main sentence of a relative construction implies the same in postposed ones.

If the definiteness of the head determines the shape of the relative pronoun, the pronoun co-occurring with the definite head will be "longer", i.e. it will contain some additional morphological element.

This last statement is based on evidence from ARABIC, COPTIC, and TANNA. Since both the definite head and the following relative pronoun are characterized by some additional marker, compared with the indefinite forms, and since the marker is, at least in ARABIC, phonologically the same for both terms, the process involved may be identified as agreement.

4.3 Given the semantic and syntactic correspondence of relative clauses and some adjectives, it would not be surprising to find that adjectives, too, agree in definiteness. This is in fact the case in ARABIC, where the definite article co-occurs with the adjective if the modified noun is definite; in GERMAN and ICELANDIC, where the adjective has "weak" and "strong" forms, the alternation is correlated with the definiteness of the head. That the "weak" and "strong" alternation has indeed the same function as article marking is suggested by the further fact that a language has either two articles in a noun-adjective construction or one article and the above-mentioned adjective alternation, but apparently never both. Further, whereas both the adjective and the noun bear a definiteness marking but no element of the relative clause in a few languages (e.g. ICELANDIC), no language has been found where definiteness agreement between the head and the relative clause would not have implied agreement for the adjective and its head as well.

[10] In LATIN, PORTUGUESE, SPANISH, and ALBANIAN, subjunctive is also used if the clause expresses opinion or unaccomplished action.

The term "agreement" has been used here rather tentatively, since some properties of definiteness agreement make it different from, say, agreement in gender. First of all, this type of agreement is strikingly rare. Second, gender agreement may operate for subject and predicate, but definiteness agreement may not. Third, it is really not quite clear what agrees with what. In the case of gender, the agreement feature is invariant throughout various syntactic processes, i.e. it is an inherent noun feature. But since the definiteness feature presumably changes during relativization, from indefinite to definite, if it were the head noun that the adjective agreed with, it would always have to be definite. Since this is not generally the case,[11] it is more likely that the adjective agrees with the definiteness feature of the whole noun phrase it belongs to.

Certain adjectives are restricted as to what article they may co-occur with. Taking the term "adjective" in its most general sense of meaning simply a non-sentential noun modifier, such restrictions of course hold for demonstrative and possessive adjectives discussed in the previous section. Furthermore, various quantifiers belong here. Although the question of quantifiers has not been explored in any detail whatever, some interesting facts can be mentioned, such as that "every" never occurs with a noun that has a preprefix in LUGANDA; it is a strong adjective in ICELANDIC, but it may still co-occur with a noun that has a definite article. "All" generally turns out to be anomalous with respect to order: in ALBANIAN, SYRIAN ARABIC, ANCIENT GREEK, HUNGARIAN, MAASAI, and MACEDONIAN it may either precede or follow the noun, and it never co-occurs with an indefinite article but stands either without an article or with a definite article. In SAHIDIC COPTIC the relative pronoun that follows "all" is sometimes definite, sometimes indefinite; in BOHAIRIC COPTIC it is always definite. In HUNGARIAN the preposed "all" takes the indefinite verb ending as an object, whereas postposed it is definite.[12] "Each" and "both" are always definite modifiers whenever a language can somehow indicate this.

---

[11] As a matter of fact, ALBANIAN, BONTOK, GAYO, KANGEAN, MADURESE, and TAGALOG have a so-called "definite article" which always appears with the adjective regardless of the definiteness status of the head.

[12] Arnauld and Lancelot claim that tous determines the noun; but their notion of determination differs from ours. In the Grammaire, such words as ce, quelque, plusieurs, nul, aucun, sorte, espèce, genre, and various numerals are all said to determine the noun inasmuch as they specify its extension, as opposed to a noun phrase that contains none of these. The underlying idea behind defining determination in terms of specification was to show that, although Vaugelas' law according to which all relative heads must have a definite article in French was false, there was indeed a general restriction on relative heads, i.e. they had to be determined, in the above sense. Lancelot, in his Logique, has a different view of determination based on the notion of restriction rather than specification--a view that is more akin to the one taken in this paper. (For more on this, see Donzé 1967.)

Let us now take the term "adjective" in a somewhat more specific sense, designating modifiers that can be derived from relative clauses. We have already seen that, besides demonstratives and possessive pronouns, "same" always makes a noun definite. "Other" appears to be wholly unlimited in this regard; in at least one language, MAASAI, the only clear minimal nonpronominal pairs with spect to definiteness are those noun phrases co-occurring with "other"; one has the prefix and means "the other (thing)", the second has no prefix and means "another (thing)". We may set up the following hypothesis: obligatorily marked definiteness contrast for any nonpronominal noun phrase implies such a marked contrast for noun phrases including "other". Two further classes of adjectives found to be significant from the viewpoint of definiteness marking are superlatives and ordinals. The only restriction on superlatives is that they must co-occur with a definiteness marker (in ALBANIAN, GREEK, in the SCANDINAVIAN languages, and in well-known WESTERN EUROPEAN languages). Ordinals are derived by means of some definite element such as the definite article in such languages as ALBANIAN, EFATE, FIJIAN, MALAY, NIAS, SAMOAN. In EGYPTIAN and SYRIAN ARABIC, no article occurs with ordinals. In a number of INDOEUROPEAN languages and in HUNGARIAN, ordinals tend to occur only with the definite article. In ATHAPASCAN the definite marker is used with "first". In CLASSICAL ARABIC ordinals always seem to occur with definite nouns. In ICELANDIC ordinals are always "weak", that is, they have only one series of endings, the one descriptive adjectives have when they modify indefinite nouns (except for "second" which is always "strong"). BULGARIAN ordinals have long and short forms except for "seventh" and "eighth" which have only the long, i.e. definite form.

4.4 No general rules have been found governing the definiteness marking of appositions, mainly because the concept itself is cloudy. In ICELANDIC appositive adjectives are said to be "strong". In EGYPTIAN ARABIC they are definite. BULGARIAN and COPTIC appellatives, if they precede the noun, have no definite article, but do if they follow. In OLD PERSIAN, too, the appellative may go before or after the proper name, but if it has the definite article it must follow.

NUMERATION

5.1 In a recent paper, Perlmutter (1968) has shown that the ENGLISH indefinite article must be described as an unstressed cardinal. He adduces as evidence that "a" contrasts with cardinals, that the collocational properties of "a" resemble those of cardinals; and that the phonological shape itself has actually evolved from "one".

The phonological similarity of the indefinite article to the cardinal "one" is a widespread phenomenon. Excepting KWAKIUTL, where some demonstrative is said to mark indefinite nouns, CLASSICAL ARABIC, GIPSY, and TONKAWA, where a suffix, apparently unrelated to the numeral system, signals indefiniteness, MARANAO, where a topic marker may mark indefiniteness, and CHINOOK, where it is marked by the neuter gender, in all languages explored the indefiniteness marker is identical or similar to the numeral "one".

For some languages, this marker is obligatory; for others, grammars claim "no indefinite article", but cardinal "one" "could sometimes be used in that function." Basing our statement on the tenuous support of our finding no grammar which stated that "one" could not be so used, and on the commonsense principle of trying to make maximally general statements, we tentatively conclude that cardinal "one" is an optional indefiniteness marker in all languages.

If the indefinite article is "one", we might expect that non-numerable nouns cannot take the indefinite article. This has been found to be true; although names of materials and of abstracta can be freely definitized, they cannot usually occur with the indefinite article. COPTIC is an exception, where the indefinite article derived from cardinal "one" is used with names of materials that are unnumeratable.

If the indefinite article is cardinal "one" and the one difference between the two is stress, it would follow that noun phrases containing other cardinals are also indefinite. In other words, whereas "this boat" can never be indefinite and "boat" must have an indefinite article to be indefinite, we predict that "five boats" is indefinite as it is. However, if numerated nouns behaved like definite noun phrases, it would refute this claim.

In no language of the sample could a numerated noun take an indefinite article derived from "one". In CLASSICAL ARABIC if a numerated noun construction takes a relative clause, the relative pronoun is one that generally occurs after indefinite heads. In HUNGARIAN transitive verbs have indefinite endings if their object is a numerated (but nondefinitized) noun phrase.

Furthermore, in languages with both definite and indefinite articles, the plural noun without the article, unless used generically, is indefinite; in other words, the plural of the indefinite article is zero. This fact is easily explained if we regard the indefinite article the same as "one". Some languages, such as PORTUGUESE, have a plural indefinite article derived from the stem "one"; others, like COPTIC, use suppletive forms.

Besides its function as an indefiniteness marker, the indefinite article is also used with generic nouns. If it is simply one of the cardinals we would expect that noun phrases containing other cardinals can also be used generically. This turns out to be true, at least for some languages ("Two people could do it").

Thus, several properties of the indefinite article can be understood if we assume that it is a numeral. Our reasoning so far is based on the following: "All numerals are indefinite or generic; therefore, a (stressless) numeral 'one' is indefinite or generic." Rather than accepting it as an axiom, a natural question would be, Why are numerals indefinite?

According to our framework, indefinite noun phrases underlie first mentionings of (simple) nouns. Since there appear to be valid reasons for regarding numerals as derived rather than underlying, we are now searching for a way to derive numerals from first mentionings, to explain their indefiniteness. As hinted on

page 65, just such a derivation is suggested by Sanders who contends that while underlying (count) nouns are always singular and "one" is a predictable feature, numerated and pluralized nouns should be derived from their conjunction. Therefore, the fact that numerated noun phrases and those containing the indefinite article are indefinite proves to be derivable from our definitization rule and is thus taken to be explained.

ON THE NOTION "ARTICLE"

6.1 The gist of the preceding discussion has been the following: nouns, i.e. pronominalizable entities, are either generic or specific; if specific and not marked for definiteness in the first place, they are subject to a definitization rule provided they occur twice in a discourse referring to the same object. Definiteness is a pronominal category; it is manifestly present in pronouns and is distinctive in the third person singular (and may be for demonstratives and question pronouns) and redundant otherwise. There are three presumably universal definiteness markers: a restrictive relative clause which underlies all (common) definite noun phrases but which becomes optional on the surface, due to an optional deletion rule; a demonstrative or third person pronoun; and some adjective meaning "same". Obligatory and nonuniversal definiteness markers include the third person pronoun, the definite article, order restrictions, and tone patterns. Optional and presumably universal indefiniteness markers are "some" and "other" and, as the noun is numerable, the cardinal "one"; the indefinite article is an obligatory and nonuniversal marker.

We now turn to the following question: What is an article and how might articles be described in a maximally general way such that the difference between a language that has these elements and one that does not can be characterized in a precise and meaningful manner.

Throughout this paper we have taken the terms "indefinite article" and "definite article" to designate markers obligatorily used with (some) first and second mentionings of noun phrases, respectively; we have taken them to mean what most grammars tacitly assume. Although we have talked about definite and indefinite noun phrases and definite and indefinite articles, it must have become clear during the discussion that articles are neither necessary nor sufficient for a noun phrase to have one or the other value of the feature. The main question now concerns, however, not use of these adjectives in association with the term "article" but rather the term "article" itself. This term of course is fairly helpful in informal discussions; but the question is whether subsuming the indefinite and the definite article under the same category in a grammar can be justified.

6.2 A theoretical term is presumably justified if it refers to some set of things constituting a natural class. If articles form a natural class, it must be possible to show that, more often than not, they behave the same way. Next, the behavior of articles will be tested for similarity with respect to the following criteria:

1. Phonological shape
2. Order
3. Agreement
4. Distribution with respect to noun classes
5. Distribution with respect to language types

1. Phonological shape - Assuming that definite and indefinite articles are subclasses of the same category, one might expect them to have some phonological similarity, at least in some languages, say, a common stem. No such language has been encountered, and the possible phonological differences between them are apparently unlimited. In BAMBARA the definite article (and for some nouns the indefinite article) is a tone pattern, while the indefinite article for other nouns also includes a segmental marker. As for segmental articles, they seem to share only two properties: brevity and stresslessness. No article with more than two syllables has been found. These features, however, elusive as the first one is, are shared by members of a great many other word classes, such as pronouns and prepositions, and thus do not count as evidence for the relationship of the two articles.

Grammars of several languages claim that only one article means both "the" and "a certain". Such languages (ATJEH, BONTOK, FIJIAN, FUTUNA, HAIDA, KUPANG, LUGANDA, MAASAI(?), PANGKUMU, RUKUMBU, SAMOAN) might have a "specific" article whose absence means "generic" rather than a "definite" and an "indefinite" article, both of which are specific and both of which may mark genericity; but this question, like a great many others raised in this paper, can be answered only after more informant work, or at least a thorough study of actual texts, rather than mere consultation of descriptions.

2. Order - If both the definite and the indefinite article of a particular language either preceded or followed the noun, the ordering rule would be more comprehensive, since it could refer to "article" instead of to two different categories. A sample of 14 languages gives the following results:

a. Definite article preposed, indefinite article postposed
1 (CLASSICAL ARABIC)
b. Indefinite article preposed, definite article postposed
2 (ALBANIAN, BULGARIAN)
c. Both articles postposed 1 (TONKAWA)
d. Both articles preposed 10 (COPTIC, ENGLISH, FRENCH, GERMAN, HUNGARIAN, ITALIAN, MARANAO, PORTUGUESE, SPANISH, YIDDISH)

If results are similar in a larger sample, it would show that for at least one type of rule it might be useful to represent articles as belonging to a single category.

3. Agreement - Provided that articles have the same associated categories, such as gender and number, the rule specifying agreement between noun and article would be more general if articles constituted a single formal category. Few data are decisive at this point, but in some examples the definite article represents categories that the indefinite article does not (e.g. gender in

COPTIC) or the indefinite article has associated categories that the definite article does not (such as CLASSICAL ARABIC and SAMNANI; in the latter the indefinite article and the word meaning "other" are the only elements of the language that have gender contrast).[13]

4. Distribution with respect to noun classes - As we have seen, any noun that is pronominalizable may take a definite article, but of these, only those that are numeratable may take an indefinite article. This is true even if articles are genericity markers. In other words, indefinite articles are more limited in distribution than definite articles, since two condition, not just one, determine the former.

5. Distribution with respect to language types - It is of course not true that languages which have definite articles also have indefinite ones, and vice versa, although this is the case for the majority of languages (for exceptions see Appendix II, page 93). Accordingly, as we would expect, articles can apparently be both gained and lost in the history of a language. AZTEC used to have a definite but no indefinite article; it now has only an indefinite article. PERSIAN seems to have lost the definite article. The history of the ROMANCE languages, of HEBREW, and of EGYPTIAN illustrates the acquisition of both definite and indefinite articles. An interesting question, of course, is whether presence or absence of one or the other article can be correlated with some other feature of the language, in that particular historical stage. During a rather cursory investigation, no typological criteria were found in only those languages which have or which lack obligatory definiteness or indefiniteness markers; all that can be said now is that if there are such properties, they must be sought among features of language that change relatively fast, rather than among more pervasive one. On the other hand, languages that have articles can sometimes be subsumed under genetic or areal cover terms.

6.3 Thus, articles prove to be independent entities when compared with each other, and we have failed to find any rules, except perhaps ordering rules, which would benefit from putting definite and indefinite articles into one category: conditions determining their distribution and phonological setup vary too greatly. The question that we raise now, therefore, is whether definite and indefinite articles are unique categories or are they subsumable under some other linguistic categories?

A partial answer has already been given in discussing various bits of evidence. In particular, the last section shows that (almost) all properties of an indefinite article can be derived from properties of numerals. The indefinite article marks count nouns, indefinite or generic, just as all other cardinals do; it usually has no plural, and neither does the numeral "one". Further pertinent criteria testing the relationship of cardinals and indefinite articles might be ordering and associated categories. As for order, the only two languages where the obligatory indefiniteness marker was found postposed were CLASSICAL ARABIC and TONKAWA; though the ordering of number with respect to the noun in TONKAWA is unknown to me, the number precedes the noun in ARABIC.

[13] Professor Charles A. Ferguson called this to my attention.

Notice, however, that in both cases the indefinite articles are among the few which are apparently unrelated to the numeral system. Otherwise an obligatory indefiniteness marker, if any, precedes the noun and so do the cardinals. Evidence concerning the respective agreement properties of the indefinite article and cardinal "one" show no exception to their being identical, except for scattered plurals of the indefinite article.

Thus, the difference between cardinal "one" and the indefinite article is basically one of stress; if it is indeed true that all language have stressless "one" as an optional indefiniteness marker, the difference between languages which have and which lack an indefinite article is the difference between an obligatory and an optional rule. The deficiency of this treatment is that it does not account for obligatory indefinite markers which originate outside the numeral system. Such languages are few in number and hard to account for.

On the other hand, we have also seen some evidence that definite articles are in some respects similar to pronouns. Let us briefly summarize the relevant points. First, for a noun phrase to contain a definite article, it must be pronominalizable: the conditions for reducing a noun phrase to a pronoun or for adding a definite article appear to be exactly the same. Second, whether or not they have definite articles, languages always have an optional definiteness marker, like what accompanies definite objects in certain languages, and these too turn out to be pronouns. Third, the definite article, in most cases, looks like a pronoun, third person singular or demonstrative. Exceptions are MARANAO, where a topic marker appears to have the appropriate function, and ALBANIAN, BIERIAN, HIDATSA, and WELSH, where the definite article is apparently not homophonous with any other element of the language.

The question about agreement properties and ordering should be raised again. It is a general rule that definite articles never have more associated categories than at least one pronoun of that language, and they often have more than the nouns. In some languages pronouns are the only (or additional) elements that denote the number of nouns by co-occurring with them (a personal pronoun in FIJIAN, a demonstrative in AMHARIC); in other languages only pronouns and definite articles denote number (BENGALI and TUNICA).

As for order, definite articles are mainly preposed, like indefinite articles (see Appendix, points 1. and 2.). Schwyzer (1936) proposed that suffixed definite articles have developed from preposed articles of adjectives that followed the noun. This does not seem to be so--if it were, suffixed articles would imply free or noun-plus-adjective order, which is not true for ICELANDIC, unless at some earlier stage. Another reason for believing that the prenominal is the more basic position for definite articles is that it makes more sense, provided definite articles are indeed pronouns. In that case, the ordering could be explained in terms of a genus-before-species ordering rule suggested by Sanders (1967, section 3.1). Independent evidence for the existence of such a rule, he points out, might come from phrases such as "someone good" but not "*good someone"; "men and women" rather than "*women and men" or by the ordering of object nouns and object pronouns (the former following, the latter preceding

the verb) in FRENCH. Returning to the question of the respective ordering restrictions on definite articles and pronouns, we have compared demonstrative-and-noun constructions and definite-article-and-noun phrases in a sample of 26 languages (brackets are used as in Appendix II).

1. Noun + Demonstrative and Definite Article + Noun [TIRURAY], WELSH	2
2. Noun + Definite Article and Demonstrative + Noun ALBANIAN, BENGALI, BULGARIAN, MACEDONIAN	4
3. Noun + Demonstrative and Noun + Definite Article [BESEMAH], LOMA, SOMALI, TOBA	4
4. Demonstrative + Noun and Definite Article + Noun CLASSICAL and SYRIAN ARABIC, COPTIC, [EFATE], ENGLISH, FIJIAN, FRENCH, FUTUNA, GERMAN, ANCIENT GREEK, HUNGARIAN, ITALIAN, [JAVANESE], PORTUGUESE, SPANISH, TAGALOG	16

This shows that definite articles indeed tend to be ordered like demonstrative pronouns when they co-occur with nouns.

In light of such evidence and since indefinite articles are ordered, in most cases, like numbers, the observation that most definite and indefinite articles are ordered as either both preceding or both following the noun loses much of its significance. Because order was the only possible consideration in favor of a category including both the definite and the indefinite article, whereas there are a number of additional arguments for regarding the indefinite article as a number and the definite article as a pronoun, we conclude that whatever rules hold for the respective ordering restrictions on definite and indefinite articles, they should be derivable from similar rules for the respective ordering constraints on cardinals and demonstratives. That these two kinds of elements do indeed share at least some common ordering laws was pointed out in one of Greenberg's universals (1963, page 86), according to which if the descriptive adjective precedes the noun, both the numeral and the demonstrative tend to precede it also.

Thus it seems that many properties of definite articles can be best accounted for if we regard them as pronouns.[14] However, at least three points remain unexplained in terms of this suggestion. One, already mentioned, is that even if

[14] For valid or less valid reasons, various linguists have repeatedly suggested the idea that pronouns and articles are intimately related. For a recent proposal see Postal (1966). For an earlier version see Priscian (Keil, Grammatici Latini II, p. 54, quoted in Yvon 1955, p. 162): "Secundum stoicos vero quinque sunt eius partes: nomen, appellatio, verbum, pronomen sive articulus, conjunctio . . . articulis autem pronomina connumerantes finitos ea articulos vocabant, ipsos autem articulos, quibus nos caremus, infinitos articulos dicebant; vel, ut alii dicunt, articulos connumerabant pronominibus et articularia pronomina vocabant." For other, post-Stoic and pre-Postalian reasonings, see Yvon (1955).

definite articles never have more associated categories than pronouns, we would expect them always to have the same categories, never fewer as happens sometimes, or never the same but with a different value than that of an anaphoric pronoun, as we noted for GERMAN. Second, whereas definite articles may mark generic nouns, demonstratives may not. This suggests that definite articles, although generally closer phonologically to demonstratives, should be regarded as third-person pronouns rather than as demonstratives since third-person pronouns can be generic. This treatment would also provide a more coherent picture of the system of demonstratives and third-person pronouns. We can now say that a demonstrative differs from a third-person pronoun in terms of some space feature but agrees in most other properties, including that both may stand by themselves or may co-occur with a noun; if we do not regard the definite article as a third-person nondemonstrative pronoun, such a pronoun would also differ from a demonstrative, since in most languages it cannot co-occur with a noun.

6.4 In sum: language X and language Y, such that language X has no definite article and language Y has no indefinite article, differ from language X_1 and language Y_1, respectively, such that language X_1 has a definite article and language Y_1 has an indefinite article, in terms of the optionality status of a nonsemantic rule. There are universal co-occurrence restrictions on certain noun classes and articles; given familiarity with the numeral and pronominal system of a particular language, in most cases statements can be made concerning ordering, agreement properties, and phonological shape of the definite and the indefinite article. As for genericity marking in terms of articles, all we can say is that co-occurrence restrictions of articles with respect to noun types (based on the features of pronominalizability and countability) extend also to generic nouns, and that even if a language has articles, some generic nouns have no article.

APPENDIX I
INDONESIAN ARTICLES IN RELATION TO PRONOUNS

FORM	PERSONAL ARTICLE	ARTICLE	LIGATIVE	DEM. PRO.	PERS.PRO.	POSS. PRO.	REL.PRO.	INT.PRO.
si-	Bontok, Tagalog, Bisaya, Minankabau, Old & Modern Javanese, Bolongan, Nias, Atjeh, Karo, Malay, Toba, Sundanese	Atjeh, Nias	Nias	Tontemboani (siihi)	Taracan (sia 'he')		Nias, Gayo	Malay (ape 'what?' siape 'who?')
san-	Besemah, Madurese, Balinese, Kangean, Seraway, Old & Modern Javanese, Bontok, Lampong							
i-	Balinese, Howa, Dayak, Tiruray, Kambaira, Betsileo, Sangir	Ibanag, Tiruray	Bontok	Bolongan (inan)	Buginese,[1] Howa[2]	Tontemboani		
se-	Madurese, Kangean, Maanani		Madurese, Kangean					
sa-	Old Javanese, Mentaway, Malay (saṅ)							Gayo
ki-	Bolang-Mongondu, Sundanese, Balinese						Lampong (sai)	
pun-	Old/Modern Javanese, Old Macassarese, Balinese							
ra-	Old Javanese, Howa, Betsileo							
ni-	Ilocano, Balinese	Howa, Betsileo	Howa	Busang				
la-	Bimanese, Buginese							
ken-	Bontok (dat. of san-) Bisaya (dat. of si-)	Bolang-Mongondu[3]			Bimanese (3rd p. sing.)			
kai-	Bolang-Mongondu[4] Tagalog[5]							
si-/ s-	Mentaway							
tai-	Mentaway							
ti-	Ilocano			Busang				

[1] ag 'I' ivag emph. 'I' [2] aho 'I' izaho emph. 'I' [3] dat. of in- [4] dat. of ki- [5] dat.-acc. of si-

FORM	PERSONAL ARTICLE	ARTICLE	LIGATIVE	DEM. PRO.	PERS.PRO.	POSS.PRO.	REL.PRO.
ṅ-		Balinese, Baree, Old Javanese, Tontemboani	Bisaya, Buginese Bontok, Sawunese Tagalog, Toba				
aṅ-		Bisaya, Tagalog, Old Javanese					
na(ṅ)-		Bontok, Kambaira	Nias, Tagalog, Minankabau	Minankabau			Toba
in-		Bolang-Mongondu, Tontemboani*					
ikan-		Old Javanese					
nu-							Sundanese
-a		Rottinese, Old/New Macassarese, Gayo	Nias	Gayo	Tagalog+		
-itu	Malay	Besemah, Malay, Seraway		Besemah, Seraway, Malay, Tontemboani			
-e		Balinese, Buginese, Sangir		Karo		Dayak (3rd)	Buginese
-(e)de		Bimanese, Buginese					
-na			Bisaya, Toba			Kambaira	
-iya	Old Macassarese			Buginese, Old Macassarese			
-la	Basa Sangiang	Kupang					
-aya	Old Macassarese			Old Macassarese			
-lia/ -li/-la	Kupang			Malay			
-ne		Sawunese					
-i		Toba					
-enda/ -(e)ndai		Karo		Karo			

*acc., gen. of en-

+suffix in 3rd person

APPENDIX II

The following chart shows the distribution of segmental markers for first mentionings of nouns (except "other") and for non-first mentionings of nouns (except "same" and relative clause). Markers are obligatory unless + is flanked by parentheses, in which case they are optional, or by square brackets, in which case I do not know whether they are obligatory or optional. PRE and POST indicate ordering with respect to the noun. Question mark indicate there is no information on existence and/or ordering of the marker. Only the obligatory markers have been considered for languages where they are both obligatory and optional.

LANGUAGE	DEFINITE ARTICLE	INDEFINITE ARTICLE
Acholi	(+) POST	?
Albanian	+ POST	+ PRE
Aleut	(+) PRE	?
Amharic	+ POST	?
Arabic	+ PRE	Classical + POST Syrian (+) POST Egyptian (+) PRE
Aramaic	+ POST	?
Armenian	+ POST	?
Athapascan	(+) PRE	?
Atjeh	+ PRE	?
Aztec (Modern)	?	+ PRE
Baki	(+) PRE	(+) POST
Balinese	+ POST	?
Bambara	?	+ POST
Baree	[+] PRE	[+] POST
Basque	[+] POST	[+] POST
Bengali	[+] POST	?
Besemah	[+] POST	?
Betsileo	[+] PRE	?
Bierian	[+] PRE	[+] PRE
Bimanese	[+] POST	?
Bisaya	[+] PRE	?
Bolang-Mongondu	[+] PRE	?
Bolongan	[+] PRE	?
Bontok	+ PRE	?
Buginese	[+] POST	?
Bulgarian	+ POST	(+) PRE

LANGUAGE	DEFINITE ARTICLE	INDEFINITE ARTICLE
Celtic (British, Irish)	+ PRE	?
Chiluba	(+) PRE	?
Chinese	(+) PRE	?
Chinook	(+) PRE	(+) PRE
Chitimacha	(+) PRE	(+) PRE
Coptic	+ PRE	+ PRE
Dakota	[+] ?	[+] ?
Danish	+ POST	?
Dayak	+ POST	?
Efate	[+] PRE	[+] PRE
English	+ PRE	+ PRE
Eskimo	(+) ?	?
Fijian	+ PRE	[+] PRE
French	+ PRE	+ PRE
Fulani	(+) POST	?
Futuna	+ PRE	(+) PRE
Gayo	(+) POST	?
German	+ PRE	+ PRE
Gipsy	?	[+] POST
Greek (Ancient)	+ PRE	(+) POST
Haida	(+) PRE	?
Hebrew	+ PRE	?
Hidatsa	(+) POST	(+) POST
Howa	[+] PRE	+ ?
Hungarian	+ PRE	+ PRE
Ibanag	+ ?	?
Icelandic	+ POST	(+) PRE
Ilocano	[+] ?	?
Italian	+ PRE	?

LANGUAGE	DEFINITE ARTICLE	INDEFINITE ARTICLE
Javanese	[+] PRE	?
Kambaira	[+] PRE	?
Karo	[+] ?	?
Kupang	+ POST	[+] POST
Kwakiutl	+ ?	+ ?
Lampong	(+) ?	?
Loma	+ POST	[+] POST
Luganda	(+) PRE	?
Maasai	+ PRE	?
Macassarese	[+] POST	?
Macedonian	+ POST	?
Mae	(+) PRE	?
Malagasy	+ PRE	?
Malay	(+) POST	?
Maloese	(+) PRE	?
Maori	+ ?	?
Maranao	+ PRE	+ PRE
Masaretic	[+] ?	(+) ?
Mentaway	(+) ?	?
Nabaloi	+ PRE	?
Nias	+ PRE	?
Pangkumu	[+] PRE	?
Persian (Old)	(+) PRE	?
Ponca	[+] POST	(+) POST
Portuguese	+ PRE	+ PRE

LANGUAGE	DEFINITE ARTICLE	INDEFINITE ARTICLE
Romanian	+ POST	+ PRE
Rottinese	[+] POST	+ ?
Rotuman	?	[+] ?
Russian	(+) PRE	(+) ?
Samoan	[+] PRE	[+] PRE
Sangir	[+] POST	?
Santo	+ PRE	(+) PRE
Sawunese	[+] POST	?
Seraway	[+] ?	?
Somali	+ POST	?
Spanish	+ PRE	+ PRE
Sundanese	?	(+) ?
Tagalog	+ PRE	[+] PRE
Tanna	[+] PRE	[+] ?
Taracan	[+] ?	?
Tettum	[+] PRE	[+] PRE
Tigre	[+] PRE	?
Tiruray	[+] PRE	?
Tlingit	[+] ?	?
Toba	+ POST	?
Tonkawa	+ POST	+ POST
Tontemboan	[+] ?	?
Tunica	[+]	?
Welsh	+ PRE	?
Wolof	(+)	?
Yiddish	+ PRE	+ PRE

1. Summary

MARKER	OBL	OPT	OBL/OPT
indefinite	17	14	14
postposed	3	4	5
preposed	13	7	6
definite	47	22	35
postposed	16	5	10
preposed	25	15	16

2. Coexistence of indefiniteness and definiteness markers

↓ INDEF / DEF →	OBLIGATORY	OPTIONAL	OBL/OPT ?
OBL	Albanian, Coptic, English, Classical Arabic, French, German, Italian, Spanish, Hungarian, Maranao, Portuguese, Romanian, Tonkawa, Yiddish 14		Rottinese 1
OPT	Egyptian & Syrian Arabic, Bulgarian, Futuna, Ancient Greek, Icelandic, Santo 7	Baki, Chinook, Chitimacha, Hidatsa, Russian 5	Masaretic, Ponca, Tettum 3
OBL / OPT	Fijian, Kupang, Loma, Tagalog 4		Baree, Basque, Bierian, Dakota, Efate, Samoan, Tanna 7

BIBLIOGRAPHY

Some items of interest focused on determination are referred to below. For a more complete list of relevant papers and books, see General Bibliography.

ARABIC	Cowell chap. 18 & 19	1964	Fischer	1959
BAMBARA	Bird	1968		
CHINESE	Annear	1965		
COPTIC	Bell	1968	Till	1927
ENGLISH	Annear	1965	Moore & Shopen	1967
	Bach	1968	Perlmutter	1968
	Baker	1966	Postal	1966
	Barton	1962	Robbins	1961
	Brewer	1966	Shopen, Diller, Partee	1967
	Hill	1950	Smith	1961
	Kuroda	1968	Smith	1964
	Lakoff	1968b	Wolfe	1968
FRENCH	Donzé	1967	Guillaume	1919
GREEK	Kiekers	1921		
HITTITE	Sturtevant	1930		
INDONESIAN	Brandstetter	1913		
LATIN	Trager	1932		
PERSIAN	Kent	1944		
ROMANIAN	Löfstedt	1942		
SERBOCROATIAN	Reichenkron	1966-		
TIGRE	Palmer	1961		
GENERAL	Brugmann	1904	Schwyzer	1936
	Heinimann	1965	Trubetzkoy	1966
	Reiler	1967	Yvon	1955

Charles Ferguson

WORKING PAPERS ON LANGUAGE UNIVERSALS

Number One

November 1969

Language Universals Project
Committee on Linguistics
Stanford University
Stanford, California